

Assessing the relative importance of human and spatial pressures on non-native plant establishment in urban forests using citizen science

Katherine Duchesneau¹, Lisa Derickx¹, Pedro M. Antunes¹

¹ Department of Biology, Algoma University, Sault Ste-Marie, Ontario, Canada, P6A 2G4

Corresponding author: Pedro M. Antunes (pedro.antunes@algomau.ca)

Academic editor: R. Bustamante | Received 3 March 2021 | Accepted 17 April 2021 | Published 10 May 2021

Citation: Duchesneau K, Derickx L, Antunes PM (2021) Assessing the relative importance of human and spatial pressures on non-native plant establishment in urban forests using citizen science. NeoBiota 65: 1–21. <https://doi.org/10.3897/neobiota.65.65415>

Abstract

Since 2007, more people in the world live in urban than in rural areas. The development of urban areas has encroached into natural forest ecosystems, consequently increasing the ecological importance of parks and fragmented forest remnants. However, a major concern is that urban activities have rendered urban forests susceptible to non-native species incursions, making them central entry sites where non-native plant species can establish and spread. We have little understanding of what urban factors contribute to this process. Here we use data collected by citizen scientists to determine the differential impacts of spatial and urban factors on non-native plant introductions in urban forests. Using a model city, we mapped 18 urban forests within city limits, and identified all the native and non-native plants present at those sites. We then determined the relative contribution of spatial and socioeconomic variables on the richness and composition of native and non-native plant communities. We found that socioeconomic factors rather than spatial factors (e.g., urban forest area) were important modulators of overall or non-native species richness. Non-native species richness in urban forest fragments was primarily affected by residential layout, recent construction events, and nearby roads. This demonstrates that the proliferation of non-native species is inherent to urban activities and we propose that future studies replicate our approach in different cities to broaden our understanding of the spatial and social factors that modulate invasive species movement starting in urban areas.

Keywords

biological invasions, citizen science, city planning, non-native plants, urban forests

Introduction

Contemporary non-native species introductions and dispersal are intimately associated with human activity (Elton 2020) and, as we could expect, these species tend to be adapted to human-modified environments such as urban environments (van Kluenen et al. 2015). While preserving forest remnants within cities can be vastly beneficial (Alvey 2006), urban area expansion also reduces and fragments habitats, which can facilitate invasion by non-native species (Marzluff 2005; Catford et al. 2011).

The type of land use within the urban matrix represents a primary pathway for the introduction of non-native plants. For example, cities contain numerous individually managed gardens, as well as vacant unmanaged land, degraded sites and intensively managed parks where non-native species often abound. This increases the probability of their escape into urban forests (Andrén 1994), which in virtue of their disturbance, fragmentation and isolation from the original main forested area, offer, often times, suitable habitats for non-native species to establish and grow in abundance (Potgieter and Cadotte 2020). Consequently, urban forest planning and management practices should incorporate the unique attributes of urban forests and detach, to some degree, from practices specifically used in natural forests. However, a major challenge is that our understanding of invasion patterns in urban forests is lacking (Cadotte et al. 2017).

Urban forest fragments in a matrix of urbanization can, to a certain extent, be seen as akin to islands, where non-native species establishment may more readily occur due to their relative isolation and fragmentation (Davis and Glick 1978). In this context, we can formulate hypotheses to identify factors driving urban forest plant community composition. These can include, for example, the species-area relationship, which would suggest that the size of an urban forest relates to species richness (MacArthur and Wilson 1963, 1967). In this context, large urban forests can conserve higher native species diversity than smaller ones (Honnay et al. 1999), thereby being relatively more resistant to non-native species introductions than smaller ones. Conversely, edge environments are more prone to non-native species establishment; as a consequence, small urban forests can host more non-native species than their larger counterparts. Indeed, smaller natural forest fragments have previously been shown to have higher levels of invasion (Ohlemüller et al. 2006).

Spatial arrangement (e.g., distance and connectivity) among urban forests can also play a role in the susceptibility of urban forests to non-native species introductions and consequent invasions. For example, it can be hypothesized that the degree of isolation among urban forest fragments affects species richness with consequences for non-native plant richness (MacArthur and Wilson 1963, 1967; see Bastin and Thomas 1999). If this is the case, conservation measures relating to urban forest fragments and biodiversity may be informed if a distance effect among urban forests is noticeable. Furthermore, the relatively small size and proximity of urban forests to areas of high human traffic and activity makes them vulnerable to edge effects (Murcia 1995), and non-native species incursions (Ohlemüller et al. 2006). Thus, native and non-native species composition of urban forests can be influenced by both spatial variables and

attributes associated with the surrounding urban matrix (Kupfer et al. 2006). For example, non-native plant species can be abundantly brought into the urban matrix either intentionally (e.g., garden ornamentals, bird seeds) and unintentionally (e.g., via increased vectors and pathways), and thus exert a high propagule pressure on urban forests compared to native forest species that must traverse the urban matrix to reach the same forests (Nascimento et al. 2006). In fact, urban forests have more non-native species when located closer to the center of urban cores (Mack and Lonsdale 2001; Kowarik 2008). Additionally, native species within urban forests may experience greater rates of population declines the further apart they are from the nearest original and/or unfragmented forest (Gascon et al. 1999). Whether this is the result of direct and indirect interactions with non-native species or results from reproductive isolation is unknown. As such, elucidating what urban forest features contribute to non-native species richness can contribute to answer this question.

One can hypothesize that the type of land use, and the socioeconomic urban layout can be predictors of non-native species richness in urban forests. Both species richness and propagule pressure from non-native species brought into the urban matrix can vary from one neighborhood to another (Aronson et al. 2015). For example, Fan et al. (2019) studied the effects of urban landscape variables on forest community structure in Illinois, and found that while industrial, commercial and transportation land use decreased the diversity and canopy cover of trees, residential land use had a positive effect on those variables. Therefore, land use must be considered when assessing the factors determining species composition and invasion patterns in urban forests. Certainly, the relative importance of spatial factors and urban matrix factors remains unclear.

A limitation of urban studies is associated with their inherent complex spatio-temporal scales (Ohlemüller et al. 2006; Jenerette et al. 2016) and the high cost of training and deploying teams capable of conducting plant community census. A solution to this problem is the use of citizen science, which includes both a strong pedagogical component locally, and simultaneously can produce valuable datasets for a variety of urban forests in a region (Fuccillo et al. 2015). Several studies that measured the accuracy of datasets collected using citizen science consistently indicate that citizen science is an adequate tool for large-scale ecosystem surveys (Aceves-Bueno et al. 2015, 2017). Furthermore, a review by Dickenson et al. (2010) on the use of citizen science for data collection highlights the importance of combining data from separate citizen science programs to adequately monitor trends in ecology across large spatial and temporal scales.

The aim of this study is to investigate factors associated with non-native plant invasions patterns in urban forests, including spatial and land-use attributes of the urban matrix. We provide a citizen science method of data collection and an easily reproducible analysis pipeline to facilitate future studies. Using a template city, we test the hypothesis that urban forest size and relative isolation (i.e., distance to the nearest unfragmented natural forest) guide the incidence of non-native versus native plants species in urban forests. Alternatively, we hypothesize that characteristics of the urban matrix associated with layout and land-use could interfere with the distance and area effects. More specifically, we hypothesize that: (1) larger urban forests closer to the

nearest unfragmented forest have greater overall native and non-native species richness as compared to smaller and more distant urban forests, even when accounting for urban matrix characteristics such as population density and surrounding construction events through time; (2) larger urban forests have a smaller ratio of non-native to native species as compared to small ones, even when accounting for urban matrix characteristics such as population density and surrounding construction events through time. Alternatively, urban matrix characteristics may be better predictors of non-native species composition in urban forests and; (3) urban forests that occur closer together show greater similarity in species composition than those farther apart.

Methods

Study area

The study area is the city of Sault Ste. Marie, Ontario, Canada. The city lies within the Algoma District bordering the eastern shore of Lake Superior. Sault Ste. Marie has a population of approximately 75,000 (Statistics Canada 2012). The city is part of the Ontario Shield Ecozone, which is characterized by having a large portion of exposed bedrock (Crins et al. 2009). Average daily temperatures during the summer months range from 15.5–17.6 °C with 888.7mm of mean annual precipitation. Located in the center of the Great Lakes-St. Lawrence forest region, the dominant forest types comprise a mixture of both deciduous and coniferous species. Abundant tree species include sugar maple (*Acer saccharum*), yellow birch (*Betula alleghaniensis*), eastern hemlock (*Tsuga canadensis*), white and red pine (*Pinus strobus* and *P. resinosa*), white and black spruce (*Picea glauca* and *P. mariana*), and balsam fir (*Abies balsamea*) (Wake 1997). The city of Sault Ste. Marie is surrounded by relatively undisturbed forested areas which are connected to the vast boreal forests of Canada (Power and Gillis 2006). For the purpose of testing our hypotheses, we consider this to be the 'unfragmented forested area'.

Urban forest study sites

We obtained the list of study sites by identifying areas zoned by the city of Sault Ste. Marie as parks and recreational areas. We only selected undeveloped and unmanaged forested areas for inclusion in this study and excluded all other heavily managed areas such as sports fields and golf courses. A total of 18 accessible urban forests were identified ranging in size, from 2,200 to 140,5480 m² (Table 1, Figure 1). We included in our analysis a forest area that is connected to the unfragmented forest, fragment number 18 (Table 1, Figure 1), and considered it as a control for isolation by distance. We measured urbanization based on the minimum distance to a high traffic road (henceforth “distance to the closest road”), the percentage of land assigned as ‘commercial’ within a 250 m buffer of the urban forest (henceforth “commercial zoning”), the percentage of land assigned as rural within a 250 m buffer of the urban forest (henceforth

Table 1. Urban forest measurements in Sault Ste. Marie, Ontario. Site number refers to urban forest number of Figure 1.

Site number	Area (m ²)	Perimeter (m)	Distance from Forest (m)
1	2200	271	3480
2	10385	591	4572
3	4274	337	5664
4	6457	431	5299
5	10720	790	4018
6	11963	573	4870
7	11340	467	4284
8	15901	494	3925
9	14062	942	5829
10	17938	708	8496
11	34336	1491	5260
12	11588	616	4669
13	157632	1872	2291
14	315182	2800	1405
15	378112	3514	4196
16	596530	4330	145
17	1405480	6084	7877
18	1028248	5244	0

“rural zoning”), the percentage of land assigned as residential within a 250 m buffer of the urban forest (henceforth “residential zoning”), and the average year of construction within a 250 m buffer of the urban forest (henceforth “average building age”) for each urban forest. We obtained the Sault Ste. Marie zoning information from Sault Ste. Marie City Hall (Christopher Bean, GIS Coordinator).

Sampling design and data collection

We randomly distributed Modified-Whittaker sampling plots within each of the urban forest islands. The Modified-Whittaker sampling design detects greater species richness and is a more convenient sampling method than the Whittaker plot design (Ghorbani et al. 2011). It is a nested vegetation sampling design that allows sampling species richness at multiple scales and to plot species-area relationships. The design nests smaller sub-plots within one main 1000m² (20 × 50 m) plot. Sub-plots consist of one 100 m² (5 × 20 m) plot placed within the centre of the main plot, two 10m² (2 × 5 m) plots in opposite corners of the main plot, and ten 1m² (0.5 × 2 m) plots placed around the border of the main plot (Stohlgren et al. 1995).

A team of 52 citizen scientists went to each of the pre-established plots in the urban forests between July 2 and August 9, 2013 and collected data that enabled the identification of all vascular plants present to species-level. In addition, the citizen scientists counted and provided cover estimates within each of the 1m² sub-plots for each species identified. Citizen scientists used general field knowledge as well as personal field guides to help identify each plant to species. For non-native species, we provided a booklet containing descriptions of the most common invasive plants in the area. Specimens that could not be immediately identified to species were collected and tagged.

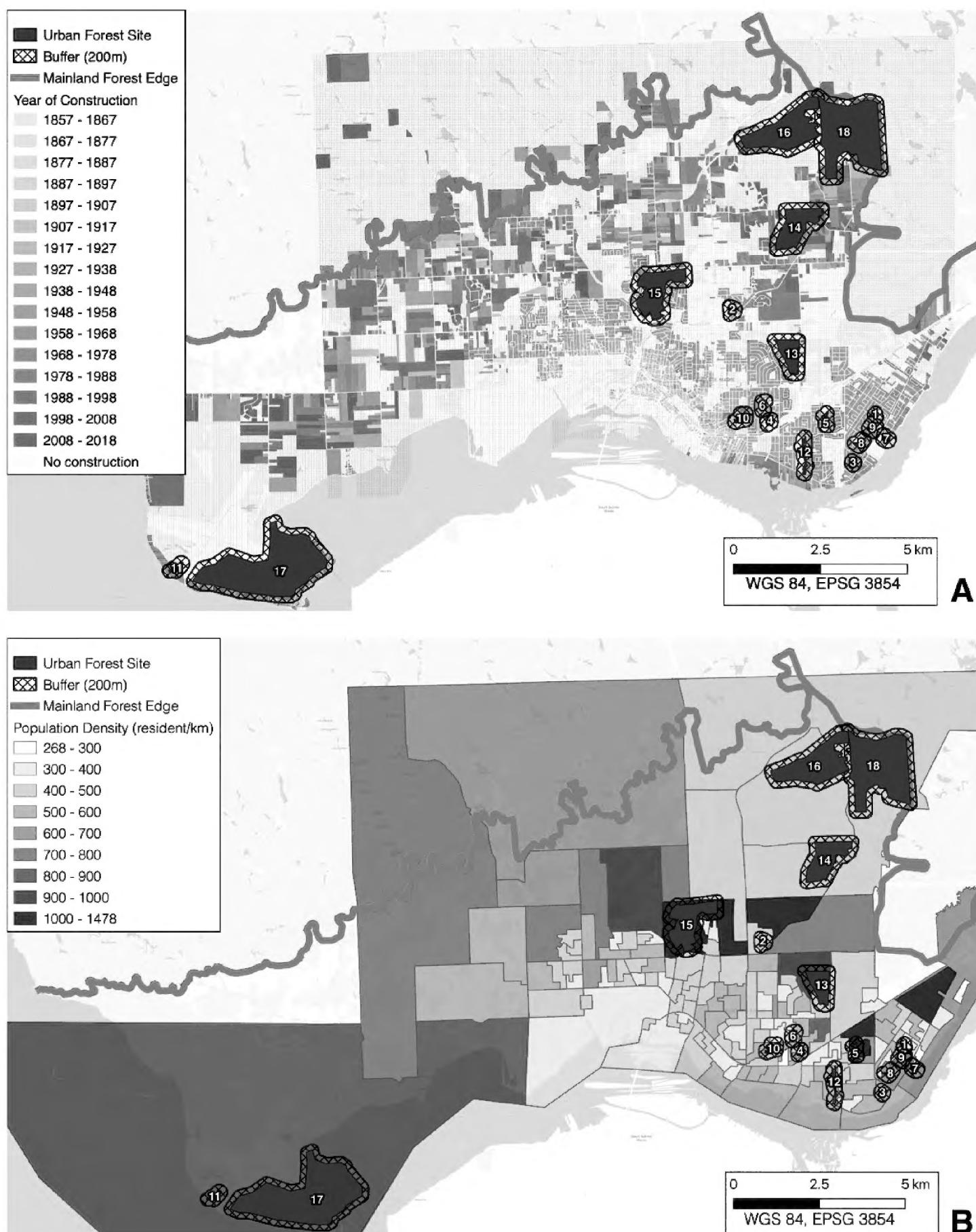


Figure 1. Map (i.e., WGS84 projection) of the 18 urban forests in Sault Ste. Marie, Ontario, Canada, in which this study is based. Both maps contain a representation of each urban forest, site 1 through 18, in green with a 250 m buffer beginning at the edge of the forested area and the unfragmented forest edge in red **A** the year of construction of all the plots with some development is indicated as a blue gradient **B** the population density of each neighborhood is indicated as a gray gradient.

We later identified these specimens with support from the scientists at the Northern Ontario Herbarium, and prepared them to be stored as part of the collection in the Algoma University Herbarium. Cover was estimated using the Braun-Blanquet cover-

abundance scale (Wikum and Shanholtzer 1978). We then created cumulative plant species lists and cover estimates for each of the 10m² and 100m² sub-plots and for the main 1000m² plot. We then assigned species a native or non-native status according to the PLANTS Database (USDA, NRCS 2018).

Geometry calculations

We calculated all geometry features in QGIS (QGIS Development Team 2019). Using Google Earth Pro satellite images, we produced polygons around the urban islands and identified the edge of the larger forest surrounding the city of Sault Ste. Marie (henceforth the “unfragmented” forest). All distances were calculated from a map projected with the coordinate reference system utm16N. We measured area (m²) and length (m) of the edge around (henceforth the perimeter) the urban forest polygons. Using the NNJoin pluggin version 3.1.2, we measured the distance between each urban forest from edge to edge and the shortest distance of each urban forest to the “unfragmented” forest line.

To quantify urban landscape use, we added the location of high traffic roads, the year of construction of each plot, and zoning information for plots to our map. We computed the distance between each urban forest and the closest high traffic road. The city zones include rural, environment/natural, mining, park, residential, and commercial land use. To determine impacts of adjacent urban factors on urban forest composition, we used a 250 m buffer, which starts at the edge of each forest and ends 250 m within the urban matrix, around each urban forest. We calculated the percentage designated to each zone in the buffer area, and the average year of construction of plots in the buffer area (Table 1, Figure 1).

Statistical analysis

We used R version 3.5.1 (R Core Team 2018) to perform all statistical analyses. We produced two types of variables for our analyses. The first type of variables described spatial arrangement of the urban forests and included area, perimeter, and distance to the “unfragmented” forest. The second type of variables describe the urban landscape surrounding the urban forests and included the distance to the closest road, the commercial zoning, the rural zoning, the residential zoning, and average building age. To avoid collinearity among predictor variables, we conducted pair-wise Pearson’s correlation tests and kept all variables with a correlation coefficient below 0.7 and above -0.7 (Dorman et al. 2013). When faced with choosing between two variables with a correlation coefficient above 0.7 or below -0.7, we kept the variable that was most ecologically meaningful. Area and perimeter of urban forests were positively correlated ($R = 0.96$). We kept area in the final model because of the relevance of species-area relationship in ecology (e.g., Connor and McCoy 1979), which was central to our hypotheses. Rural zoning was negatively correlated to residential zoning and positively correlated with area. Again, because of the importance of area in our analysis, we chose to remove rural zoning and keep residential zoning as a variable in the final model.

Since we did not have *a priori* hypotheses about which variables (i.e., spatial and socio-economic) could better explain species composition, we performed multi-model inference to rank candidate models using Akaike's information criterion corrected for a large number of predictors (AICc). We used this technique to determine which independent variable could better explain the variance in species richness, non-native to native species ratio, non-native species richness, and native species richness. We tested all response variables for normality using the Shapiro-Wilks test. We nested all possible combinations of both the spatial and the urban landscape variables to produce a set of candidate models. To assess variable significance, we calculated the weighted model average of all candidate models within the 95% confidence set of models (sum of model weights > 0.95) (Burnham and Anderson 2004). We considered variables to be relevant when the confidence intervals of the variables in the averaged model did not overlap zero. The estimates of averaged models cannot reliably measure the effect size of a variable (Cade 2015). Instead, we reported the adjusted R-Squared of a model including only the relevant variables and the variance explained by each variable through hierarchical partitioning using *hier.part* from the *hier.part* package (Nally and Walsh 2004), which provides a reliable assessment of the strength of the correlation between the dependent and the independent variable (Chevan and Sutherland 1991). We also analyzed the data using 'classical inference', and found that in both analyses similar conclusions were supported by the results (Table 2).

To determine if spatial and landscape variables influenced species composition, we performed a redundancy analysis (rda) and partitioned the variance between spatial and urban landscape variables. We used Hellinger transformed matrices of species composition as the response variables of the rda with the function *rda* from the package *VEGAN* (Oksanen et al. 2018). We included area, area to perimeter ratio, distance to the "unfragmented" forest, distance to roads, all zoning components, and the average year of construction surrounding the urban forest as response variables. None of the variables had a variance inflation factor higher than 3. The significance of the redundancy analysis was tested using *ordiR2step*, from *VEGAN*, with 1000 permutations. The significance of each axis and then each term was assessed similarly using *anova.caa*, from *VEGAN*. We repeated this analysis with only the non-native and the native species to see if the same patterns were found in all the groups.

To test for the relationship between the distance among urban forests and their similarity in community assembly, we used the function *mantel.correlog* from the package *VEGAN*, with Pearson's correlation coefficient and 1000 permutations. We provided the program with a distance matrix of the distance among the urban forests and a similarity matrix composed of species abundance for each urban forest. We calculated the break point between distance classes for the mantel correlogram using the Sturges equation. We repeated this analysis by replacing the similarity matrix with a matrix containing only the non-native species, and then only with the native species to see if different factors affected each group. Both the R script of the analysis and the dataframe of used variables are available as supplementary files (see Supple material 1: 'analysis' for R script with the analysis, and Supple material 2: 'data_csv' for the dataframe of used variables).

Table 2. List of the estimates for the most parsimonious model for each variable.

Variables	AICc			F- statistics*	
	Estimate	Lower confidence interval	Upper confidence interval	Estimate	P-value
Kept as part of the model predicting species richness					
Distance from the “unfragmented” forest	0.004	-0.0008	0.0091	0.002	0.542
Kept as part of the model predicting the ratio of native to exotic species					
Residential zoning	0.0036	0.0019	0.005	0.02	0.0001
Average year of construction	0.0077	0.0032	0.0125	0.02	0.239
Distance to the closest road	0	0	0.0001	0.0002	0.0029
Kept as part of the model predicting exotic species richness					
Distance to the closest road	0.0014	0.0004	0.0026	0.002	0.01
Residential zoning	0.1743	0.0751	0.2665	0.17	0.007

* Additional information on the statistical methods is available as supplementary material (Supple material 4: F-statistics).

Results

Hypothesis 1: Relationship between urban forest size, distance, and species richness

We found a total of 142 plant species across the urban forests surveyed. Of these species, 36 were non-native and 106 were native. Each urban forest plot contained an average of 16.5 species (± 7.64). On average, 24.19 % (± 21.07) of plants in each plot were non-native, with an average of 4.33 non-native species (± 3.61) and 12.17 native species (± 5.25) per Whitaker plot. The most abundant species were *Maianthemum canadense* (present in 72.22% of plots), *Rubus pubescens* (38.89%), *Fraxinus americana* (33.33%), *Rubus idaeus* (44.44%) and *Acer saccharum* (38.89%). Of these, only *Rubus idaeus* is listed as non-native (see Supple material 3: ‘species data’).

A total of 64 candidate models were compared to find the best way to predict overall species richness from both spatial and landscape variables. Of the candidate models, 36 were kept as part of the confidence set of the average model, including the Null model (Table 2). The most parsimonious model (with the lowest AICc value) only included distance from the “unfragmented” forest (Estimate (E) = 0.0040) (Figure 2). However, when considering the confidence interval of the effect of distance to the “unfragmented” forest (Confidence interval (CI) = -0.0008, 0.0091), it could not explain species richness. The other variables, including the distance to the closest road (CI = -0.0019, 0.0061), commercial zoning (CI = -1.3975, 0.4731), residential zoning (CI = -0.0246, 0.0484), and average building age (CI = -0.0572, 1.6543), were also not considered as being meaningful by the confidence set model average.

Hypothesis 2: Relationship between urban forest size, distance, urban factors, and native and non-native species richness

We produced three separate groups of candidate models to test which variables affected the ratio of non-native to native species, as well as the non-native and native species, independently. Each response variable produced a total of 64 candidate models. When

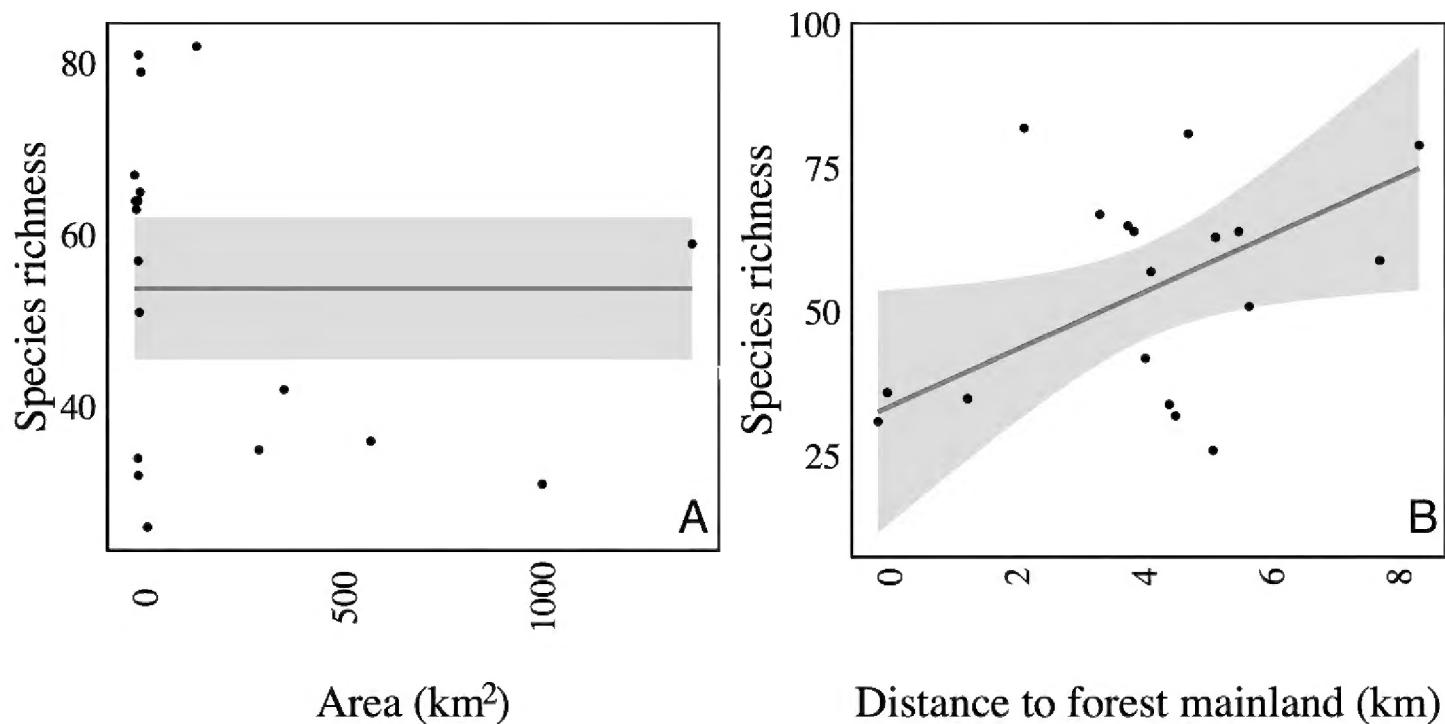


Figure 2. Each graph represents the linear relationship between urban forest species richness and a variable while controlling for all other variables. Species richness was not related to the area of urban forests (A), and it was also not related to distance between the urban forest and the unfragmented forest (B).

analyzing the effect of spatial and urban landscape factors on the ratio of non-native to native species, seven models were kept as part of the confidence set of the average model. Area and distance to the unfragmented forested area were not included in the most parsimonious models (Figure 3A, B). The most parsimonious model included residential zoning ($E = 0.0036$), the average year of construction ($E = 0.0077$) (Figure 3C), and the distance to the closest road ($E = 0.0000$) (Figure 3D). The average model from the confidence interval set of models provided similar results; residential zoning ($CI = 0.0019, 0.0050$), average year of construction ($CI = 0.0032, 0.0125$), and distance to the closest road ($CI = 0.0000, 0.0001$) had confidence intervals that showed a positive effect on the non-native to native species ratio. In contrast, commercial zoning ($CI = -0.4368, 0.0572$), area ($CI = -0.0000, 0.0000$), and distance from the “unfragmented” forest ($CI = -0.0022, 0.0024$) did not add to the predictive ability of the model. According to the R-squared adjusted, together, residential zoning, average year of construction, and distance to the closest road explained 69.35% of the variation in the non-native to native species ratio. When we assigned importance to the relevant variables through hierarchical partitioning, we found that residential zoning was the most important variable, with an importance level (I) of 56.48%. However, the distance to the closest road and the average building age were still moderately important (I = 21.88% and 21.63% respectively).

When analyzing the effect of spatial and urban landscape variables on non-native species alone, 18 models were kept as part of the confidence interval set for the model average. The most parsimonious model included a positive effect of the distance to the closest road ($E = 0.0014$) and residential zoning ($E = 0.1743$) on the number of non-native species. Similarly, only the distance to the closest road ($CI = 0.0004,$

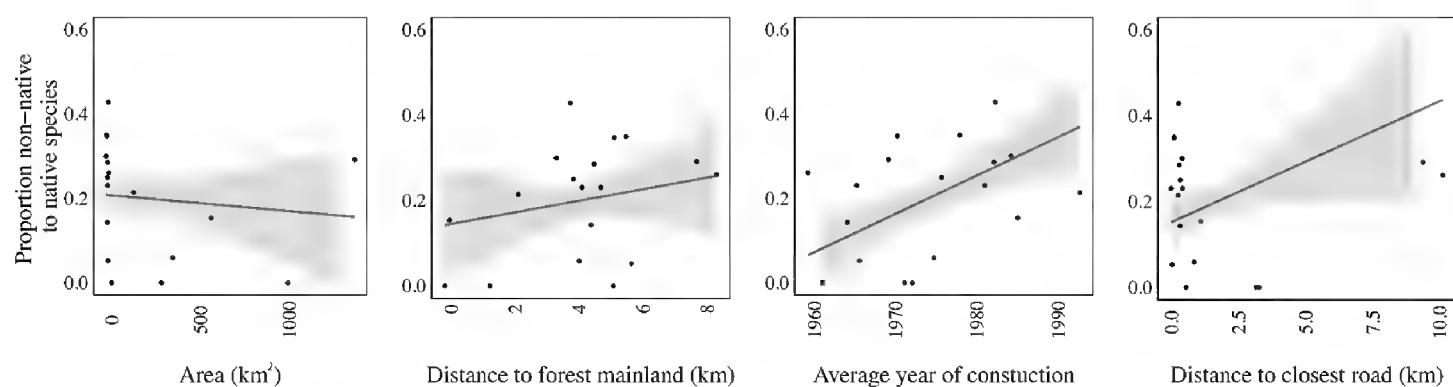


Figure 3. Linear relationships between the proportion of non-native relative to native species in urban forests and each variable while controlling for variables in the other panels **A** the proportion of non-native to native species was not related to the area of urban islands, calculated in meters squared **B** the proportion of non-native to native species was not related to distance between the urban island and the unfragmented forest, calculated in meters **C** the proportion of non-native to native species increased with newer constructions **D** the proportion of non-native to native species increased with distance from the closest road, calculated in meters.

0.0026) and residential zoning (CI = 0.0751, 0.2665) had confidence intervals that did not include zero in the model average. Despite the effect of the average year of construction (CI = -0.0662, 0.5207) on the non-native to native species ratio, it did not affect non-native species richness. Together, the distance to the closest road and residential zoning explained 49.43% of the variation in non-native species richness, and residential zoning (I = 72.11%) was more important than the distance to the closest road (I = 27.89%). Commercial zoning (CI = -0.4368, 0.0572), area (CI = -0.0000, 0.0000), and distance from the “unfragmented” forest (CI = -0.0022, 0.0024) did not add to the predictive ability of the model. When analyzing the effect of spatial and urban landscape variables on native species alone, 31 models were kept as part of the confidence interval set for the model average. The most parsimonious model was the null model. The model average had no variable with confidence intervals that did not cross zero. Despite the effect of residential zoning (CI = -0.2936, 0.2477), average year of construction (CI = -0.5293, 1.1767), and distance to the closest road (CI = -0.0018, 0.0040) on the non-native to native species ratio, they did not affect native species richness. Commercial zoning (CI = -1.0031, 0.4744), and the spatial variables area (CI = -0.0000, 0.0000) and distance from the “unfragmented” forest (CI = -0.0015, 0.0066) were also not considered as meaningful.

We tested the effect of space and urbanization on the overall species composition, the non-native species composition, and the native species composition of the urban forests by partitioning the variance and through model selection of redundancy analysis using permutation tests. The full model could not adequately describe the variance in species composition ($F = 0.9821$, $P = 0.564$). Instead, the model that best described the overall species composition was the null model. When only the native species were modeled, we found that none of the variables could explain the patterns of native species composition. Similarly, non-native species composition was not modeled by the set of space and urbanization variables.

Hypothesis 3: Relationship between proximity of urban forests and species composition

To determine if the proximity of the urban forests influenced their overall, non-native, or native species composition we performed mantel correlogram tests with distance classes calculated using Sturge's equation. According to the mantel correlogram analysis, distance could not predict overall species composition ($r = 0.032$, $P = 0.372$). However, urban forests that were in the first distance class, or, in other words, close together, were similar in species composition ($P = 0.046$) (Figure 4). When separated by native/non-native status, there was a change in how distance between urban forests affected the similarity in community structure. We found that native species showed no significant spatial correlation in structure ($r = 0.001$, $P = 0.439$) (Figure 4), but non-native species had a significant negative linear correlation ($r = 0.313$, $P = 0.001$). In fact, urban forests that were closer together were more likely to have similar non-native species ($P = 0.001$) and those further apart were more likely to be composed by different non-native species ($P = 0.006$). However, the pattern breaks at the farthest distance class (Figure 4).

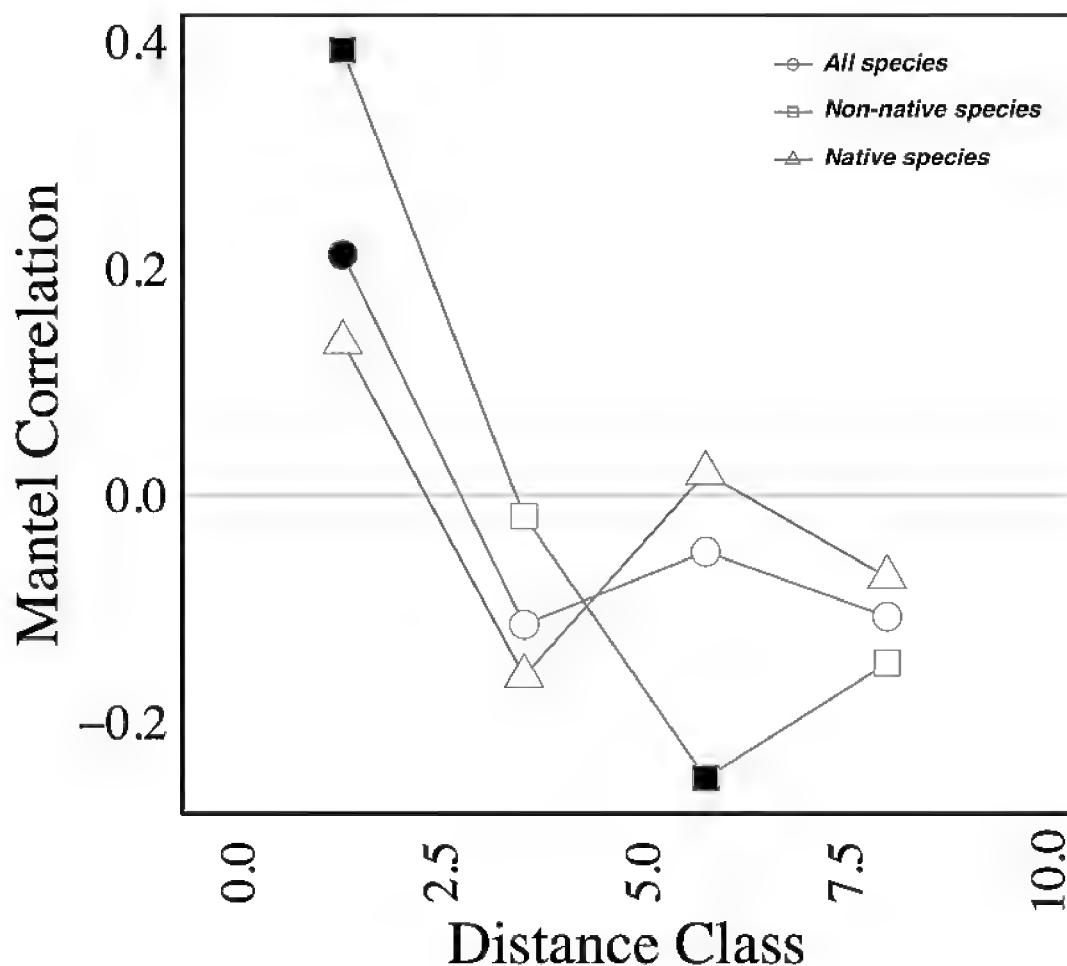


Figure 4. Mantel's correlogram defaulting to Pearson's correlation coefficient with 1000 permutations. We calculated the distance classes using Sturge's equation. The correlation of distance between urban forests on the similarity in composition of species was tested in four distance classes. Points filled with white represent distances where species composition was not related to distance and solid points represent distances where species composition was related to distance. The purple line connecting circles is the representation of the analysis for all the species. The blue line connecting triangles represent the analysis for native species. The red line joining squares represent the analysis for non-native species.

Discussion

Our approach using a novel citizen science method of data collection and analysis pipeline enabled identifying whether spatial and/or land-use attributes of the urban matrix were associated with non-native plant occurrence patterns in urban forests. First, it is important to note that the native plants species recorded in our urban forests were consistent with those present in communities of the Great Lakes-St. Lawrence forests of Ontario. The most common native species found in the urban forests, namely *Fraxinus americana*, *Acer saccharum*, and *Maianthemum canadense*, are representatives of the core community of forests in the Great Lakes-St. Lawrence region according to Canada's National Forest Inventory. Similarly, the dominant non-native invasive species reported by the citizen scientists in this study, including *Rubus idaeus* and *Alliaria petiolata*, are consistent with reports for this area (e.g., Invasive Species Awareness Program – <http://www.invadingspecies.com/>). Second, we found that spatial variables did not adequately predict the overall species richness and composition of forest fragments. Both urban forest area and distance to the nearest unfragmented non-urban forest had no effect on plant species richness, suggesting that the urbanization does not always completely isolate the urban forest fragments within the city matrix.

The lack of spatial signal raises the hypotheses that: 1) compared to more densely populated cities, the type of urban development based on detached houses with gardens, which is typical of our model city, may contribute to buffer reproductive isolation among urban forests, and; 2) our urban forests are too recent for ecological effects to have emerged (<100 years). Consistent with this perspective, forest fragments that were adjacent to the relatively less disturbed non-urban forest adjacent to the city showed no sign of hosting more native species relative to the forest fragments imbedded within the city. In contrast, species composition among urban forests responded to residential land-use, construction events, and proximity of roads, suggesting that landscaping and residential planning could be main drivers of non-native species introductions and, eventually, invasion. When considering native and non-native species separately, the proportion of land zoned for residential use surrounding urban forest fragments, recent construction events, measured by the average age of infrastructures near the urban forests, and possibly the distance between these forests and the closest road were factors correlated with an increase in non-native to native species ratio. Additionally, reductions in species composition similarity among urban forests with increasing distance from each other, particularly for non-native species, further indicates that these sites are unlikely to be completely isolated by the urban matrix around them and could be responding to local urbanization factors instead. As such, we conclude that spatial variables, at least in some cases, can be poor predictors of species community richness and non-native species community composition and, instead, we propose that emphasis should be placed on qualities of the urban matrix to determine urban forest non-native community composition.

The urban matrix and urban forest communities

Our original intent was primarily focused on spatial rather than land use characteristics that meaningfully predict plant community richness and composition. However, the urban matrix variables clearly served to account for factors that, together, could have a larger effect on community composition than the size and spatial arrangement of urban forests. Consequently, we propose that, going forward, studies to predict non-native species community assembly in other urban centers should incorporate characteristics of the urban matrix that are important in urban forest design and management.

While we found compelling evidence that the species richness of urban forests was exclusively driven by parameters relating to the urban matrix, their effect was restricted to the richness of non-native species. These results are congruent with previous studies indicating that human activities (Davis 2009), and particularly the trade of non-native plants in urban areas (Reichard and White 2001; Tartaglia et al 2018), are a major vector for non-native species introductions. The observed increase in non-native species present in the forest fragments near residential areas reinforces the notion that urban forests are likely suitable habitats for these species to escape into (Tartaglia et al. 2018). Additionally, in these residential areas, construction events near urban forests, measured as the average construction year, is a factor positively correlated with non-native species richness. There is a well-established relationship between disturbance events and invasion in the literature (Marvier et al. 2004; Didham et al. 2007; Foxcroft et al. 2011); as such, our results are unsurprising when we think of construction activities as a major potential source of non-native species. However, we found that as the distance from the main roads increased, so did the number of non-native species in forest fragments. We were surprised by these results, as several previous studies have linked road use to non-native invasive species richness (Gelbard and Belnap 2003; Von Der Lippe and Kowarik 2007; Flory and Clay 2009). However, the most common non-native species present in our sites were *Rubus idaeus* and *Alliaria petiolata*, which are primarily dispersed by animals; perhaps distance from roads could be associated with greater animal dispersal. Additionally, the incongruence in our results and the literature could be an indicator of historical human development or land-use variables. Again, although these results are not the primary findings of this research, they indicate potentially important characteristics to consider as indicators of urbanization.

This study adds to the body of knowledge on the importance of considering socio-economic factors when analyzing the diversity and species composition of urban landscapes (Hope et al. 2006; González-Moreno et al. 2013; Godefroid and Ricotta 2018; Fan et al. 2019). In fact, ignoring qualities of the urban matrix has been flagged as a limitation of solely using spatial factors; for example, in island biogeography (Laurence 2008). Island Biogeography Theory does not consider the fundamental differences between the ocean, which is largely inhospitable to island species, and the urban matrix which has repeatedly been shown to host sustainable populations of herbaceous species in yards and empty lots (Johnson et al. 2018). Qualities of the

urban matrix itself can be used to determine patterns of species composition in urban forests. For example, the proportion of sealed surface in the urban matrix, associated with the degree of urbanization, can be used to predict patterns of species richness in urban forests (Malkinson et al. 2018). Our findings that residential land use (but not commercial use) is associated with an increase in non-native species emphasize the importance of considering variation within the urbanized landscape. Additionally, taking into consideration the urban forest's historical and geographic context when considering patterns of species community assembly is essential. Habitat loss and fragmentation tends to occur in non-random patterns because certain types of habitats, for example areas of low and constant elevation, are better suited to human activity. Thus, these areas are developed more quickly, which attracts further development and loss of habitat, in turn increasing the rate of species loss (Seabloom et al. 2002). As such, it is not completely surprising that overall species patterns cannot be predicted by area and distance relationships alone.

Application to other urban centers

This study constitutes a first step towards understanding how distance between urban forests and their area affects species composition and patterns of non-native plant invasions. Even though in cities, such as the one used in this study, connectivity between urban forests and proximity to major uninhabited forests may determine the low predictive capacity of spatial attributes alone, that may not be the case in larger urbanized centers, particularly those where habitat fragmentation is high. We know from previous studies that patterns of species composition in urban forests are dependent on both anthropogenic and ecosystem factors at local and regional scales (Ohlemüller et al. 2006; Jenerette et al. 2016). For example, the bioregional context of a city determines the overall species composition of urban forests as well as trends in non-native species composition. Urban forests in the northern regions of North America tend to have less non-native tree species compared to those in the southern regions of the continent, primarily due to the limitations imposed by minimum winter temperatures on trees selected by humans for landscaping purposes (Jenerette et al. 2016). In this context, this study indicates that northern cities, which are not as fragmented or surrounded by fragmented areas as southern cities, can be relatively more protected from biodiversity losses due to non-native plant species introductions. As such, the approach used here can serve as a template to determine larger scale patterns of species composition in future studies, at least across cities in Eastern North America. While this undertaking might require many additional resources and a high degree of coordination, we believe it can be highly effective in the context of invasive species prevention and monitoring. We propose that future studies replicate our methods in other cities to ascertain whether our findings are widely applicable. Additionally, we strongly advocate the use of citizen science as a method of data collection to maximize resources and increase public awareness and knowledge of non-native plants in cities.

Acknowledgments

This research was conducted in Robinson-Huron Treaty territory and the traditional territory of the Anishnaabeg, specifically the Garden River and Batchewana First Nations, as well as Métis People. The work was funded through grants from the Ontario Trillium Foundation and Ontario Ministry of Natural Resources and Canada Research Chairs awarded to P.M. Antunes. The authors are especially grateful to the team of 52 volunteers who contributed a total of 248 hours to the project as well as to Christopher Bean (City of Sault Ste. Marie) and the staff at Northern Ontario herbarium who helped us with the taxonomic identification of some specimens.

References

Aceves-Bueno E, Adeleye AS, Bradley D, Brandt WT, Callery P, Feraud M, Garner KL, Gentry R, Huang Y, McCullough I, Pearlman I (2015) Citizen science as an approach for overcoming insufficient monitoring and inadequate stakeholder buy-in in adaptive management: criteria and evidence. *Ecosystems* 18: 493–506. <https://doi.org/10.1007/s10021-015-9842-4>

Alvey AA (2006) Promoting and preserving biodiversity in the urban forest. *Urban Forestry and Urban Greening* 5: 195–201. <https://doi.org/10.1016/j.ufug.2006.09.003>

Aceves-Bueno E, Adeleye AS, Feraud M, Huang Y, Tao M, Yang Y, Anderson SE (2017) The accuracy of citizen science data: a quantitative review. *The Bulletin of the Ecological Society of America* 98: 278–290. <https://doi.org/10.1002/bes2.1336>

Andrén H (1994) Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: A review. *Oikos* 71: 355–366. <https://doi.org/10.2307/3545823>

Aronson MF, Handel SN, La Puma IP, Clemants SE (2015) Urbanization promotes non-native woody species and diverse plant assemblages in the New York metropolitan region. *Urban Ecosystems* 18: 31–45. <https://doi.org/10.1007/s11252-014-0382-z>

Bastin L, Thomas CD (1999) The distribution of plant species in urban vegetation fragments. *Landscape Ecology* 14: 493–507. <https://doi.org/10.1023/A:1008036207944>

Burnham KP, Anderson DR (2004) Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods and Research* 33: 261–304. <https://doi.org/10.1177/0049124104268644>

Cade BS (2015) Model averaging and muddled multimodel inferences. *Ecology* 96: 2370–2382. <https://doi.org/10.1890/14-1639.1>

Cadotte MW, Yasui SLE, Livingstone S, MacIvor JS (2017) Are urban systems beneficial, detrimental, or indifferent for biological invasion? *Biological Invasions* 19: 3489–3503. <https://doi.org/10.1007/s10530-017-1586-y>

Catford JA, Vesk PA, White MD, Wintle BA (2011) Hotspots of plant invasion predicted by propagule pressure and ecosystem characteristics. *Diversity and Distributions* 17: 1099–110. <https://doi.org/10.1111/j.1472-4642.2011.00794.x>

Chevan A, Sutherland M (1991) Hierarchical partitioning. *The American Statistician* 45: 90–96. <https://doi.org/10.1080/00031305.1991.10475776>

Crins WJ, Gray PA, Uhlig PWC, Wester MC (2009) The ecosystems of Ontario. Part 1: Eco-zones and ecoregions. Ontario Ministry of Natural Resources. Peterborough, ON: Inventory, Monitoring and Assessment Section. Gen. Tech. Rep. SIB TER IMA TR-01.

Connor EF, McCoy ED (1979) The statistics and biology of the species-area relationship. *The American Naturalist*. 113: 791–833. <https://doi.org/10.1086/283438>

Davis MA (2009) Invasion Biology. Oxford University Press on Demand.

Davis AM, Glick TF (1978) Urban ecosystems and island biogeography. *Environmental Conservation* 5: 299–304. <https://doi.org/10.1017/S037689290000638X>

Dickinson J, Zuckerberg B, Bonter D (2010) Citizen science as an ecological research tool: challenges and benefits. *Annual Review of Ecology, Evolution, and Systematics* 41: 149–172. <https://doi.org/10.1146/annurev-ecolsys-102209-144636>

Didham RK, Tylianakis JM, Gemmell NJ, Ewers RM (2007) Interactive effects of habitat modification and species invasion on native species decline. *Trends in Ecology and Evolution* 22: 489–496. <https://doi.org/10.1016/j.tree.2007.07.001>

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B, Lafourcade B, Leitão PJ, Münkemüller T, McClean C, Osborne PE, Reineking B, Schröder B, Skidmore AK, Zurell D, Lautenbach S (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36: 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>

Elton CS (2020) The ecology of invasions by animals and plants. Springer Nature. <https://doi.org/10.1007/978-3-030-34721-5>

Fan C, Johnston M, Darling L, Scott L, Liao FH (2019) Land use and socio-economic determinants of urban forest structure and diversity. *Landscape and Urban Planning* 181: 10–21. <https://doi.org/10.1016/j.landurbplan.2018.09.012>

Flory SL, Clay K (2009) Effects of roads and forest successional age on experimental plant invasions. *Biological Conservation* 142: 2531–2537. <https://doi.org/10.1016/j.biocon.2009.05.024>

Foxcroft LC, Pickett STA, Cadenasso ML (2011) Expanding the conceptual frameworks of plant invasion ecology. *Perspectives in Plant Ecology, Evolution and Systematics* 13: 89–100. <https://doi.org/10.1016/j.ppees.2011.03.004>

Fuccillo KK, Crimmins TM, de Rivera, CE, Elder TS (2015) Assessing accuracy in citizen science-based plant phenology monitoring. *International Journal of Biometeorology* 59: 917–926. <https://doi.org/10.1007/s00484-014-0892-7>

Gascon C, Lovejoy TE, Bierregaard Jr RO, Malcolm JR, Stouffer PC, Vasconcelos HL, Laurance WF, Zimmerman B, Tocher M, Borges S (1999) Matrix habitat and species richness in tropical forest remnants. *Biological Conservation* 91: 223–229. [https://doi.org/10.1016/S0006-3207\(99\)00080-4](https://doi.org/10.1016/S0006-3207(99)00080-4)

Gelbard JL, Belnap J (2003) Roads as Conduits for Exotic Plant Invasions in a Semiarid Landscape. *Conservation Biology* 17: 420–432. <https://doi.org/10.1046/j.1523-1739.2003.01408.x>

Ghorbani J, Taya A, Shokri M, Naseri HR (2011) Comparison of Whittaker and modified-Whittaker plots to estimate species richness in semi-arid grasslands and shrublands. *Desert* 16: 17–22.

Godefroid S, Ricotta C (2018) Alien plant species do have a clear preference for different land uses within urban environments. *Urban Ecosystems* 21: 1189–1198. <https://doi.org/10.1007/s11252-018-0792-4>

González-Moreno P, Pino J, Carreras D, Basnou C, Fernández-Rebollar I, Vilà M (2013) Quantifying the landscape influence on plant invasions in Mediterranean coastal habitats. *Landscape Ecology* 28: 891–903. <https://doi.org/10.1007/s10980-013-9857-1>

Honnay O, Endels P, Vereecken H, Hermy M (1999) The role of patch area and habitat diversity in explaining native plant species richness in disturbed suburban forest patches in northern Belgium. *Diversity and Distributions* 5: 129–141. <https://doi.org/10.1046/j.1472-4642.1999.00047.x>

Hope D, Gries C, Zhu W, Fagan WF, Redman CL, Grimm NB, Nelson AL, Martin C, Kinzig A (2003) Socioeconomics drive urban plant diversity. *Proceedings of the National Academy of Sciences* 100: 8788–8792. <https://doi.org/10.1073/pnas.1537557100>

Invading Species Awareness Program (2020) Invading Species Awareness Program. Invasive Plants. (n.d.). <http://www.invadingspecies.com/plants/> [Accessed June 22, 2020]

Jenerette GD, Harlan SL, Buyantuev A, Stefanov WL, Declet-Barreto J, Ruddell BL, Myint SW, Kaplan S, Li X (2016) Micro-scale urban surface temperatures are related to land-cover features and residential heat related health impacts in Phoenix, AZ USA. *Landscape Ecology* 4: 745–60. <https://doi.org/10.1007/s10980-015-0284-3>

Johnson AL, Borowy D, Swan CM (2018) Land use history and seed dispersal drive divergent plant community assembly patterns in urban vacant lots. *Journal of Applied Ecology* 55: 451–460. <https://doi.org/10.1111/1365-2664.12958>

Kowarik I (2008) On the role of alien species in urban flora and vegetation. In: Marzluff JM et al. (Eds) *Urban Ecology*. Springer, Boston.

Kupfer JA, Malanson GP, Franklin SB (2006) Not seeing the ocean for the islands: The mediating influence of matrix-based processes on forest fragmentation effects. *Global Ecology and Biogeography* 15: 8–20. <https://doi.org/10.1111/j.1466-822X.2006.00204.x>

Laurance WF (2008) Theory meets reality: how habitat fragmentation research has transcended island biogeographic theory. *Biological conservation* 141: 1731–44. <https://doi.org/10.1016/j.biocon.2008.05.011>

MacArthur RH, Wilson EO (1963) An equilibrium theory of insular zoogeography. *Evolution* 17: 373–387. <https://doi.org/10.1111/j.1558-5646.1963.tb03295.x>

MacArthur RH, Wilson EO (1967) The theory of island biogeography. Princeton University Press, Princeton.

Mack RN, Lonsdale WM (2001) Humans as global plant dispersers: getting more than we bargained for: current introductions of species for aesthetic purposes present the largest single challenge for predicting which plant immigrants will become future pests. *BioScience* 51: 95–102. [https://doi.org/10.1641/0006-3568\(2001\)051\[0095:HAGPDG\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0095:HAGPDG]2.0.CO;2)

Malkinson D, Kopel D, Wittenberg L (2018) From rural-urban gradients to patch–matrix frameworks: Plant diversity patterns in urban landscapes. *Landscape and Urban Planning* 169: 260–268. <https://doi.org/10.1016/j.landurbplan.2017.09.021>

Marvier M, Kareiva P, Neubert MG (2004) Habitat destruction, fragmentation, and disturbance promote invasion by habitat generalists in a multispecies metapopulation. *Risk Analysis* 24: 869–878. <https://doi.org/10.1111/j.0272-4332.2004.00485.x>

Marzluff JM (2005) Island biogeography for an urbanizing world: How extinction and colonization may determine biological diversity in human-dominated landscapes. *Urban Ecosystems* 8: 157–177. <https://doi.org/10.1007/s11252-005-4378-6>

Murcia C (1995) Edge effects in fragmented forests: Implications for conservation. *Tree* 10: 58–62. [https://doi.org/10.1016/S0169-5347\(00\)88977-6](https://doi.org/10.1016/S0169-5347(00)88977-6)

Nally RM, Walsh CJ (2004) Hierarchical partitioning public-domain software. *Biodiversity and Conservation* 13: 659–660. <https://doi.org/10.1023/B:BIOC.0000009515.11717.0b>

Nascimento HEM, Andrade ACS, Camargo JL, Laurance WF, Laurance SG, Ribeiro JEL (2006) Effects of the surrounding matrix on tree recruitment in Amazonian forest fragments. *Conservation Biology* 20: 853–860. <https://doi.org/10.1111/j.1523-1739.2006.00344.x>

Ohlemüller R, Walker S, Wilson JB (2006) Local vs regional factors as determinants of the invasibility of indigenous forest fragments by alien plant species. *Oikos* 112: 493–501. <https://doi.org/10.1111/j.0030-1299.2006.13887.x>

Oksanen JF, Blanchet G, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O’Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2019) vegan: Community Ecology Package. R package version 2.5–6. <https://CRAN.R-project.org/package=vegan>

Potgieter LJ, Cadotte MW (2020) The application of invasion frameworks to urban ecosystems. *NeoBiota* 62: 365–386. <https://doi.org/10.3897/neobiota.62.50661>

Power K, Gillis MD (2006) Canada’s forest inventory 2001 (Vol. 408). Pacific Forestry Centre.

QGIS Development Team (2019) QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>

Reichard SH, White P (2001) Horticulture as a pathway of invasive plant introductions in the United States. *BioScience* 51: 103–113. [https://doi.org/10.1641/0006-3568\(2001\)051\[0103:HAPOI\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0103:HAPOI]2.0.CO;2)

Seabloom EW, Dobson AP, Stoms DM (2002) Extinction rates under nonrandom patterns of habitat loss. *Proceedings of the National Academy of Sciences* 99: 11229–11234. <https://doi.org/10.1073/pnas.162064899>

Statistics Canada (2012) Sault Ste. Marie, Ontario (Code 3557061) and Ontario (Code 35) (Table). Census Profile. 2011 Census. Statistics Canada Catalogue no. 98-316-XWE. Ottawa. Released October 24, 2012. <http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/prof/index.cfm?Lang=E> [accessed May 8, 2013]

Stohlgren TJ, Falkner MB, Schell LD (1995) A modified-Whittaker nested vegetation sampling method. *Vegetatio* 117: 113–121. <https://doi.org/10.1007/BF00045503>

Tartaglia ES, Aronson MFJ, Raphael J (2018) Does suburban horticulture influence plant invasions in a remnant natural area? *Natural Areas Journal* 38: 259–267. <https://doi.org/10.3375/043.038.0406>

USDA, NRCS (2020) The PLANTS Database (<http://plants.usda.gov>, 25 September 2020). National Plant Data Team, Greensboro, NC 27401–4901 USA.

Van Kleunen M, Dawson W, Essl F, Pergl J, Winter M, Weber E, Kreft H, Weigelt P, Kartesz J, Nishino M, Antonova LA (2015) Global exchange and accumulation of non-native plants. *Nature* 525: 100–103. <https://doi.org/10.1038/nature14910>

Von Der Lippe M, Kowarik I (2007) Long-distance dispersal of plants by vehicles as a driver of plant invasions. *Conservation Biology* 21: 986–996. <https://doi.org/10.1111/j.1523-1739.2007.00722.x>

Wake W [Ed.] (1997) A nature guide to Ontario. University of Toronto Press, Toronto.

Wikum DA, Shanholtzer GF (1978) Application of the Braun-Blanquet cover-abundance scale for vegetation analysis in land development studies. *Environmental Management* 2: 323–329. <https://doi.org/10.1007/BF01866672>

Supplementary material 1

R script of the analysis

Authors: Katherine Duchesneau, Lisa Derickx, Pedro M. Antunes

Data type: statistical data

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neobiota.65.65415.suppl1>

Supplementary material 2

Dataframe of used variables

Authors: Katherine Duchesneau, Lisa Derickx, Pedro M. Antunes

Data type: statistical data

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neobiota.65.65415.suppl2>

Supplementary material 3

Species data

Authors: Katherine Duchesneau, Lisa Derickx, Pedro M. Antunes

Data type: species data

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neobiota.65.65415.suppl3>

Supplementary material 4

F-statistics

Authors: Katherine Duchesneau, Lisa Derickx, Pedro M. Antunes

Data type: statistical data

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neobiota.65.65415.suppl4>